README Document for OMSO2: Aura/OMI Sulfur Dioxide Level 2 Product[1]

Goddard Earth Sciences Data and Information Services Center (GES DISC)

https://disc.gsfc.nasa.gov/

NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA

Principal Investigators: Nickolay A. Krotkov, Can Li, Code 614
nickolay.a.krotkov@nasa.gov
301-614-5553

Algorithm Developer: Can Li, University of Maryland, Code 614
can.li@nasa.gov
301-614-5616

Algorithm Support: Peter J.T. Leonard, ADNET, Code 619
peter.j.leonard@nasa.gov
301-352-4659

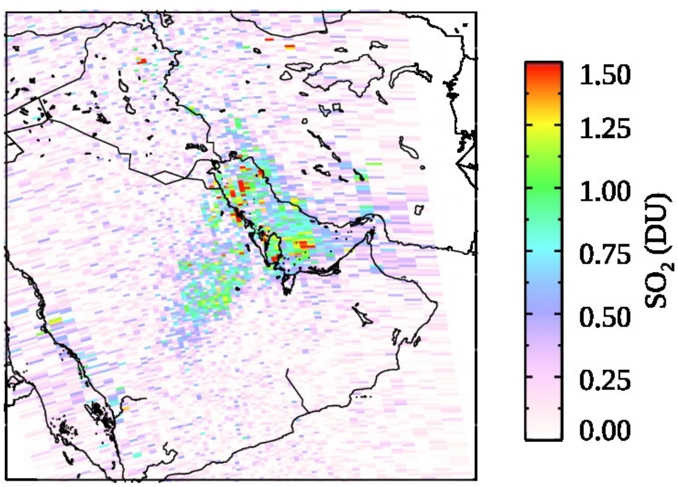

Figure 1. OMI retrieved SO_2 vertical column density in Dobson Units (1 DU = $2.69 \cdot 10^{16}$ molecules/cm^2) from orbit 5885 on August 23, 2005 showing enhanced SO_2 pollution over the Persian Gulf due to oil and gas operations in the region.

[1] The version numbers (*e.g.*, OMSO2 Version 2) in this document refer to the different versions of the OMSO2 algorithm used to generate the product. These should not be confused with "v003" in the product file names (*e.g.*, OMI-Aura_L2-OMSO2_2020m0228t0338-o83092_v003-2020m0228t155749.he5), which indicates that the product file is produced with Collection 3 OMI L1B input data.

1. **Overview**

This document describes the OMI SO_2 product (OMSO2) produced from global mode UV measurements of the Ozone Monitoring Instrument (OMI). OMI was launched on July 15, 2004 on the EOS Aura satellite, which is in a sun-synchronous ascending polar orbit with 1:45 pm local equator crossing time. The data collection started on August 17, 2004 (orbit 482) and continues to this day with only minor data gaps. The minimal volcanic SO_2 mass detectable by OMI is about two orders of magnitude smaller than the detection threshold of the legacy Total Ozone Mapping Spectrometer (TOMS) (1978-2005) [Krueger et al., 1995; Fisher et al., 2019]. OMI also enables the detection of anthropogenic SO_2 pollution in the lowest part of the atmosphere. This is due to smaller OMI footprint and the use of hyperspectral wavelengths better optimized for separating O_3 from SO_2.

Each product file, also called a data granule, covers the sunlit portion of the orbit with an approximately 2600 km wide swath. Each swath normally contains approximately 1600 viewing lines along the ground track of the satellite, with each viewing line containing 60 pixels or scenes across the satellite track. Scenes from all viewing lines with the same cross-track scene number are referred to as a row of the OMI swath. During normal operations, 14 or 15 granules are produced daily, providing fully contiguous coverage of the globe. OMSO2 product is not produced when OMI goes into the "zoom mode" for one day every 452 orbits.

Since 25 June 2007 signal suppression (anomaly) has been observed in Level 1B Earth radiance data for scenes in rows 53-54 (0-based). This anomaly is also known as the OMI row anomaly since it affects some particular rows of the CCD detector. It has since expanded to affect more rows. In SO_2 data, the row anomaly manifests itself as positive or negative stripes (discontinuity in SO_2 with cross-track viewing angle). Efforts have been made to flag the affected scenes. SO_2 data fields for scenes determined to have been influenced by the row anomaly have been assigned a fill value. Additionally, **beginning with version 2.0, a new data field named "Flag_RowAnomaly" (from OMI ozone product, OMTO3) has been added to OMSO2 product** to assist data users. More information about the OMI row anomaly can be found from KNMI.

For each OMI scene we provide 6 different estimates of the vertical column density (VCD) of SO_2 in Dobson Units (1 DU = $2.69 \cdot 10^{16}$ molecules/cm^2) obtained by making different assumptions about the vertical distribution of SO_2. In addition, we also provide one estimate of the slant column density (SCD) of SO_2 in molecules/cm^2. **The users are advised to pick one VCD estimate that suits their particular needs.** Starting from version 2.0, users interested in anthropogenic SO_2 pollution may also use their own *a priori* vertical SO_2 profiles (e.g., from a chemical transport model or from airborne measurements) and auxiliary parameters (e.g., scattering weights) provided in the OMSO2 product to produce customized VCD estimates (see section 3.5 for detailed introduction). For volcanic SO_2 VCDs, users can continue to use the SO_2 plume height (also referred to as the center of mass altitude, CMA) derived from SO_2 vertical distribution, to interpolate between the four standard VCD estimates provided for volcanic retrievals (TRL, TRM, TRU, and STL). The different estimates of SO_2 SCD and VCD in OMSO2 are:

- **SlantColumnAmountSO2**: slant column density (SCD) of SO_2 produced by spectral fitting using SO_2 cross sections. When converted to VCD with a fixed air mass factor (AMF) of 0.36, this newly introduced dataset in OMSO2 version 2.0 can be used as a continuity product for our previous version of OMI Planetary Boundary Layer (PBL) SO_2 VCD (**ColumnAmountSO2_PBL**) in OMSO2 versions 1.2 and 1.3.

- **ColumnAmountSO2**: vertical column density (VCD) of SO_2 produced by spectral fitting using updated SO_2 Jacobians that more accurately account for the effects of geometry, clouds, O_3, and surface reflectivity on OMI sensitivity as well as updated *a priori* profiles from model simulations. This new dataset introduced in OMSO2 version 2.0 is recommended for studies on air quality.

- **ColumnAmountSO2_PBL**: SO_2 VCD based on the same lookup table as ColumnAmoutnSO2, but *a priori* profiles assuming fixed mixing ratio within the planetary boundary layer (PBL, defined here as the lowest 1 km of the atmosphere) and negligible SO_2 above the PBL. It is recommended for use in estimates for SO_2 emission sources. **Note that ColumnAmountSO2_PBL in OMSO2 version 2.0 differs from the original ColumnAmountSO2_PBL dataset in the previous versions of the OMSO2 product (1.3 and earlier). Please check section 2.1 for important updates to the dataset in OMSO2 version 2.0 and later.**

- **ColumnAmountSO2_TRL**: SO_2 VCD corresponding to an assumed lower tropospheric SO_2 profile with a center of mass altitude (CMA) of 3 km. It is recommended for use in studies on degassing from volcanic sources.

- **ColumnAmountSO2_TRM:** SO_2 VCD corresponding to an assumed middle tropospheric SO_2 profile with a CMA of 8 km. It is recommended for use in studies on moderate eruptions.

- **ColumnAmountSO2_TRU**: SO_2 VCD corresponding to an assumed upper tropospheric SO_2 profile with a CMA of 13 km. It is recommended for use in studies on explosive volcanic eruptions.

- **ColumnAmountSO2_STL**: SO_2 VCD corresponding to an assumed lower stratospheric SO_2 profile with a CMA of 18 km. It is recommended for use in studies on explosive volcanic eruptions with plumes reaching the stratosphere.

The accuracy and precision of the derived SO_2 VCDs vary significantly with the SO_2 CMA and column abundance, observational geometry, ozone, surface reflectivity, and cloudiness. OMI is generally more sensitive to SO_2 above clouds and snow/ice, and less sensitive to SO_2 below clouds. Preliminary error estimates are discussed in section 4 (Data Quality Assessment).

2. Important Updates Since Initial Release

2.1 Important Updates to OMI Anthropogenic SO_2 Data (Version 2.0 and later)

In earlier versions of OMSO2 product, ColumnAmountSO2_PBL was recommended for studies and applications related to anthropogenic SO_2 pollution. In OMSO2 versions 1.2 and 1.3, ColumnAmountSO2_PBL was produced with a principal component analysis (PCA) based spectral fitting algorithm [Li et al., 2013] and fixed SO_2 Jacobians (see section 3.2 for details). The static Jacobians are a simplification and may lead to relatively large biases under certain conditions (e.g., off-nadir scenes, scenes covered by snow/ice, and cloudy scenes).

In OMSO2 version 2.0, three estimates of SO_2 abundance are provided for air quality applications: SlantColumnAmountSO2 (SCD) is the slant column density of SO_2. SCD can be converted to VCD using air mass factors (AMF). **For data users who wish to continue to use the previous ColumnAmountSO2_PBL dataset in OMSO2 versions 1.2 and 1.3, we recommend that a fixed AMF of 0.36 be used to convert SCD to VCD**, which in turn can be used as a continuity dataset in place of the original ColumnAmountSO2_PBL.

For most applications and studies related to anthropogenic pollution, we recommend the use of ColumnAmountSO2, a newly available dataset in OMSO2 version 2.0. ColumnAmountSO2 is an estimate of SO_2 VCD produced with SO_2 Jacobians from a more extensive lookup table and monthly *a priori* profiles based on model simulations. Data users are referred to section 3.4 for more information on the dataset and how it can be customized.

The new ColumnAmountSO2_PBL in OMSO2 version 2.0 differs from the original ColumnAmountSO2_PBL in versions 1.2 and 1.3 in that it is produced with SO_2 Jacobians from the same new lookup table as in ColumnAmountSO2, and different *a priori* profiles assuming well-mixed SO_2 below 1 km and negligible SO_2 above 1 km. Such a retrieval setup ensures that ColumnAmountSO2_PBL is independent from model simulations and the dataset is recommend for top-down estimates of emission sources. ColumnAmountSO2_PBL is only provided for scenes with cloud radiance fraction (CRF) < 0.5. **Please refer to Appendix A for a summary of the changes in data fields between OMSO2 Versions 1.3 and 2.0.**

2.2 Important Updates to OMI PBL SO_2 Data (Versions 1.2 and 1.3)

In OMSO2 versions 1.2 and 1.3, the SO_2 data in **ColumnAmountSO2_PBL** are produced with a retrieval algorithm based on principal component analysis (PCA) of the OMI radiance data [Li et al., 2013]. Previously the OMI PBL SO_2 data were produced using the Band Residual Difference (BRD) algorithm [Krotkov et al., 2006]. While the BRD algorithm is sensitive to SO_2 pollution in the PBL, it tends to have large noise and unphysical biases particularly at high latitudes. The PCA algorithm greatly improves the quality of OMI SO_2 retrievals and has been implemented for operational production of OMI standard SO_2 product. **In OMSO2 product version 1.2 or later, the entire PBL SO_2 data record has been reprocessed with the PCA algorithm. PBL SO_2 data users who had previously acquired the old version OMSO2 data before the public release in October 2014 are strongly encouraged to download and use OMSO2 version 1.2 or later. There is no difference in the PBL SO_2 between OMSO2 product versions 1.2 and 1.3. All SO_2 data fields ending with "BRD" are obsolete and removed starting from OMSO2 version 2.0.**

2.3 Important Updates to Version 1.3 OMI Volcanic SO_2 Data

The SO_2 data in ColumnAmountSO2_TRL, ColumnAmountSO2_TRM, and ColumnAmountSO2_STL data fields in OMSO2 product version 1.3 are produced with an extended version of the PCA algorithm [Li et al., 2017]. Previously the OMI TRL, TRM, and STL SO_2 data were produced using the linear fitting (LF) algorithm [Yang et al., 2007]. While the LF algorithm is fast and sensitive to SO_2 from volcanic eruptions, it has a tendency to underestimate large volcanic SO_2 signals due to saturation of SO_2 absorption at the strongly absorbing short UV wavelengths (< 315 nm). It also has relatively large noise and artifacts as compared with the PCA algorithm. **The entire OMI TRL, TRM, and STL SO_2 datasets in the OMSO2 product have been reprocessed with the PCA-based algorithm. Volcanic (TRL, TRM, and STL) SO_2 data users who had previously acquired OMSO2 data prior to the public release in June 2016 are strongly encouraged to download and use OMSO2 product version 1.3 or later.**

3. Algorithm Description

All OMI SO_2 data are now generated with principal component analysis (PCA)-based algorithms. Since 2014, we have released three versions of OMSO2 product based on PCA algorithms. The OMSO2 version 1.2, released in 2014, included updated PBL SO_2 VCDs produced using the original PCA algorithm with a fixed spectral fitting window and a fixed SO_2 Jacobian spectrum that are appropriate for pollution SO_2 near the surface [Li et al., 2013]. The OMSO2 version 1.3, released in 2016, had the same PBL SO_2 as in version 1.2, and updated volcanic TRL, TRM, and STL SO_2 VCDs that were produced with an extended version of the PCA algorithm, using a spectral fitting window and SO_2 Jacobian lookup table (LUT) suitable for volcanic SO_2 signals [Li et al 2017]. The new OMSO2 version 2.0, released in 2019, contains updated SO_2 data fields for both anthropogenic (SO_2, PBL SO_2, and SO_2 SCD) and volcanic (TRL, TRM, TRU, and STL SO_2) sources. The retrieval algorithms for these different versions of OMSO2 product share the same general approach and framework. Their primary differences are in the assumed *a priori* SO_2 vertical profiles and methods employed in SO_2 Jacobian calculations.

3.1 General Approach and Algorithm Framework

In the PCA-based SO_2 algorithm, we apply a principal component analysis technique to radiance data taken over a presumably SO_2-free region in the same orbital swath and row. The resulting principal components (PCs) can capture most (>99.9999%) of measurement-to-measurement spectral variation of the radiances. The PCs are ordered so that the first PC explains the most of spectral variance, the second PC explains the second most of spectral variance, and so on. The first few leading PCs are generally associated with geophysical processes including ozone absorption, surface reflectance, and rotational-Raman scattering effects (RRS, also known as the Ring effect), while the following PCs often have high-frequency features likely originating from measurement noise and artifacts such as wavelength shift and stretch. These physical processes and measurement details can cause strong interferences in SO_2 retrievals, and the PCs allow us to appropriately account for these interferences in spectral fitting. The SO_2 Jacobians represent the sensitivity of the radiances to the SO_2 VCD ($\partial N/\partial\Omega_{SO_2}$). By fitting a set of n_v PCs (v_i) along with the Jacobians to the measured Sun-normalized radiances, we can simultaneously obtain estimates of SO_2 VCD (Ω_{SO2}) and coefficients of the PCs (**ω**):

$$N(\omega, \Omega_{SO_2}) = \sum_{i=1}^{n_v} \omega_i v_i + \Omega_{SO_2} \frac{\partial N}{\partial \Omega_{SO_2}}, \qquad (1)$$

Here N is the measured N-value spectrum ($N(\lambda) = -100\times\log_{10}(I(\lambda)/I_0(\lambda)$, I and I_0 are radiance and solar irradiance at wavelength λ, respectively) for a given OMI scene. The PCA algorithm shares the same overall physics concept with the widely used Differential Optical Absorption Spectroscopy (DOAS) method, but the data-driven (vs. forward modeling) approach used to account for retrieval interferences reduces modeling uncertainties, enhances computation efficiency, and makes the PCA algorithm much less sensitive to instrument calibration issues. A more detailed discussion of the PCA algorithm can be found in Li et al., [2013] and Joiner et al., [2013].

For input data, the PCA algorithm uses OMI level 1B (L1B) radiance and irradiance data in the spectral window of 310.5-340 nm, as well as the O_3 column amount (Ω_{O3}) from the OMTO3 product [Bhartia and Wellemeyer 2002]. The spectral window includes the strong SO_2 absorption band at 310.8 nm and minimizes potential interferences due to stray light at shorter wavelengths. To better account for the orbit-to-orbit measurement artifacts and the different characteristics of the 60 rows of the OMI detector, we process data from each row of each orbit separately. Scenes (OMI pixels) with large slant column O_3 (S_{O3} > 1500 DU) or solar zenith angle (SZA > 75°**)** are expected to have relatively small signal-to-noise ratio and weak sensitivity to SO_2, and are filtered out before PCA. In addition, scenes likely affected by large volcanic SO_2 signals are also identified and filtered out. After data filtering, we first conduct PCA on the approximately 900-1300 remaining scenes for an entire row, without screening out polluted areas. Since SO_2 absorption is generally very weak outside of polluted and volcanic-affected areas, it is unlikely for the PC(s) associated with or affected by SO_2 absorption (v_{so2}) to be among the first few leading PCs. A correlation analysis between the PCs and the SO_2 Jacobians is then conducted to determine the number of PCs (n_v) to be included in the fitting. This ensures that n_v is sufficiently small to prevent the inclusion of v_{so2} and collinearity in Eq. 1, and allows reasonable initial estimates of SO_2 (Ω_{SO2_ini}) to be obtained. To maintain computational efficiency, we set an upper limit for n_v (20 for the original PBL retrievals and current volcanic SO_2 retrievals, 30 for anthropogenic retrievals in version 2.0). A second step PCA is then applied to scenes with small Ω_{SO2_ini} (within ±1.5 standard deviations for each orbit/row) to extract a new set of PCs to update Eq. 1, followed by updated retrievals of SO_2. This step is repeated twice, as the changes in the retrieved SO_2 generally become very small within two iterations. The second step PCA and retrievals are carried out separately for three segments of each row: a "tropical" region with relatively small S_{O3} or SZA, and two regions north and south of it. These regionally derived PCs more closely match the measurements and help reduce retrieval biases.

3.2 Jacobians for the Original PBL SO_2 Retrievals (Versions 1.2 and 1.3)

The SO_2 Jacobians used in the production of the ColumnAmountSO2_PBL dataset in OMSO2 versions 1.2 and 1.3 are calculated with the VLIDORT radiative transfer code [Spurr 2008]. The calculation assumes the same measurement conditions as those in the BRD algorithm. More specifically, we assume fixed surface albedo (0.05), surface pressure (1013.25 hPa), as well as fixed solar zenith angle (30°) and viewing zenith angle (0°). For SO_2, a climatological profile over the summertime eastern U.S. is used. For O_3 and temperature, the OMTO3 standard mid-latitude profiles with Ω_{O3} = 325 DU are used. This setup allows direct comparison between the new and old OMI PBL SO_2 data based on the BRD algorithm.

3.3 Implementations Details for Volcanic SO_2 Retrievals (Version 1.3)

The PCA-based volcanic (TRL, TRM, and STL) SO_2 retrieval algorithm [Li et al., 2017] is an extended version of the original PBL SO_2 PCA algorithm. The algorithm consists of two parts. The first part is essentially identical to the PBL SO_2 algorithm as described above and is used to provide input to the second part of the algorithm, including initial estimates of SO_2 VCD (Ω_{SO2_ini}) and principal components (PCs) to be used in spectral fitting. The second part of the algorithm produces more accurate estimates of volcanic SO_2 VCDs by conducting iterative spectral fitting and by using a more comprehensive lookup table for SO_2 Jacobians. We followed an approach similar to that in TOMS and OMI total column O_3 retrievals [Bhartia and Wellemeyer 2002], and used simple Lambertian equivalent reflectivity (SLER or R) derived at the surface [Ahmad et al., 2004] to implicitly account for the combined effects of aerosols, clouds, and the surface on the spectral dependence of TOA (top of the atmosphere) radiances and SO_2 Jacobians. We also neglected the effects of non-elastic rotational Raman scattering (RRS) on SO_2 Jacobians. As a result, the backscattered radiances at TOA (I) for multiple elastic Rayleigh scattering can be calculated with the following equation:

$$I = I_0(\theta_0, \theta) + I_1(\theta_0, \theta)\cos\phi + I_2(\theta_0, \theta)\cos 2\phi + \frac{R I_r(\theta_0, \theta)}{(1 - R S_b)}. \tag{2}$$

The first three terms (I_0, I_1, and I_2) on the right-hand side (RHS) of the equation represent the atmospheric component of the back-scattered radiances, while the last term represents the surface component. θ_0, θ, and ϕ stand for solar zenith angle (SZA), viewing zenith angle (VZA), and relative azimuth angle (RAA), respectively. Taking the partial derivative with respect to Ω_{SO2}, we can obtain the following equation used for determination of SO_2 Jacobians:

$$\frac{\partial I}{\partial \Omega_{so2}} = \frac{\partial I_0(\theta_0, \theta)}{\partial \Omega_{so2}} + \frac{\partial I_1(\theta_0, \theta)}{\partial \Omega_{so2}}\cos\phi + \frac{\partial I_2(\theta_0, \theta)}{\partial \Omega_{so2}}\cos 2\phi + \frac{R}{(1 - R S_b)}\frac{\partial I_r(\theta_0, \theta)}{\partial \Omega_{so2}} + \frac{R^2 I_r(\theta_0, \theta)}{(1 - R S_b)^2}\frac{\partial S_b}{\partial \Omega_{so2}} \tag{3}$$

Using VLIDORT, we built a set of pre-computed multi-dimensional SO_2 Jacobian lookup tables, with eight SZA nodes (0-81°), eight VZA nodes (0-80°), and 15 SO_2 nodes (0-1000 DU), for each of the 21 standard O_3 climatology profiles used in OMI total O_3 retrievals [Bhartia and Wellemeyer 2002]. This was done separately for four different prescribed SO_2 profiles (TRL, TRM, TRU, and STL) at 0.05 nm spectral resolution for the spectral range of 311-342 nm.

For a given OMI scene with SZA = θ_0, VZA = θ, RAA = ϕ, initial estimate of SO_2 VCD Ω_{SO2_ini}, and O_3 column amount, Ω_{O3}, the algorithm first determines SLER (R) at 342.5, 354.1, 367.04 nm, where contributions from gaseous absorption and non-elastic RRS processes are minimal, and then extrapolates $R(\lambda)$ at these wavelengths to shorter wavelengths (λ=310-340 nm) using a second-degree polynomial function. The algorithm then determines the SO_2 Jacobian spectrum for the scene by interpolating Jacobians calculated using eq. 3 for two O_3 (corresponding to two different O_3 climatology profiles), two SO_2, two SZA, and two VZA nodes that bracket Ω_{O3}, Ω_{SO2_ini}, θ_0, and θ, respectively. Details on how the SO_2 Jacobians are calculated are given in Li et al. [2017].

The SO_2 Jacobian spectrum and the principal components from the first part of the algorithm are fit to the measured N values in the nominal spectral fitting window of 313-340 nm to produce an updated estimate of SO_2 VCD (Ω_{SO2_step1}) for the scene Ω_{SO2_ini}. If the difference between Ω_{SO2_ini} and Ω_{SO2_step1} is greater than 0.1 DU or 1% for scenes with SO_2 > 100 DU, Ω_{SO2_step1} is used as input to the lookup table to update the SO_2 Jacobian spectrum. The iterations continue until the results converge or the number of iterations exceeds 15. In each iteration step, the left edge of the actual fitting window is determined by locating the wavelength with the largest SO_2 Jacobian within 313-340 nm (up to 326.5 nm or approximately the mid-point of the nominal fitting window). All wavelengths shorter than this wavelength are excluded in the spectral fitting. This approach allows the interpolation error due to signal saturation at short UV wavelengths to be minimized.

3.4 Updates to Volcanic SO_2 Retrievals in OMSO2 Version 2.0

The volcanic SO_2 datasets in OMSO2 version 2.0 are produced using an algorithm similar to that in version 1.3 with the following updates:

1) Implementation of a new volcanic SO_2 flagging scheme that helps to detect OMI scenes with large volcanic SO_2 signals and to exclude them in PCA analysis. This scheme utilizes ozone retrieval residuals from two wavelength pairs (313 and 314 nm, and 314 and 315 nm). The residuals represent the differences between the measured and calculated radiances at different wavelengths in OMI ozone retrievals assuming that no SO_2 exists in the scene. Under most conditions when there is little SO_2 in the atmosphere, the residuals are similar between, for example, 313 and 314 nm. In the case of large volcanic eruptions, however, the residuals at 313 nm are much greater due to stronger SO_2 absorption at this wavelength. Tests with OMI data have shown that the scheme can effectively detect OMI scenes with ~5 DU of SO_2 in the stratosphere.

2) Use of the new OMI total column ozone data from OMTO3 version 9 as input in the calculations of SO_2 Jacobians.

3) Introduction of an additional SO_2 VCD dataset that assumes an upper tropospheric SO_2 *a priori* profile with a CMA of 13 km (ColumnAmountSO2_TRU).

3.5 Updates in the New Anthropogenic SO_2 Retrievals (OMSO2 Version 2.0)

While the anthropogenic SO_2 retrieval algorithm in OMSO2 version 2.0 follows the same general framework as in the original PBL SO_2 algorithm (versions 1.2 and 1.3), several updates have been implemented in the new version, including:

1) Updated data filtering. In the new OMI anthropogenic SO_2 algorithm, the same volcanic SO_2 flagging scheme as in the version 2.0 volcanic SO_2 retrieval algorithm has been implemented (see section 3.4). This helps to exclude OMI scenes affected by large volcanic eruptions in the PCA analysis and reduces the impacts of those scenes on the derived principal components and the ensuing spectral fitting. Additionally, instead of filtering out scenes based on slant column O_3 (as in OMSO2 versions 1.2 and 1.3), a simpler data filter is now used to exclude scenes with large solar zenith angle (SZA > 75°) in anthropogenic SO_2 retrievals.

2) Updated spectral fitting for areas affected by south Atlantic anomaly (SAA). The updated scheme examines the spectral fitting residual at each wavelength within 310.5-340 nm for each OMI scene located within the SAA affected area (0-45°S, 100°W-5°E). If a given scene within the area has wavelengths with relatively large fitting residuals (beyond ±0.2 *N*-value), these wavelengths are excluded in a second step spectral fitting that produces the final estimated SO_2 VCD for the scene. Tests have shown that this updated scheme effectively reduces retrieval noise over the SAA-affected area, while at the same time has negligible impacts on the retrieved SO_2 over known sources in the region.

3) Updated SO_2 Jacobian/air mass factor calculations. A table lookup approach has been implemented for SO_2 Jacobian/air mass factor (AMF) for each individual OMI scene. Scattering weight or vertically resolved box-AMF, defined as:

$$m(z) = \frac{\partial ln(I)}{\partial \tau_{SO2}(z)}, \quad (4)$$

represents the sensitivity of OMI measured TOA radiances (I) to changes in SO_2 loading (in terms of SO_2 optical thickness, τ_{SO2}) at different altitudes (z, represented by different pressure levels in the Jacobian lookup table). Scattering weight at a given wavelength λ depends on several factors including O_3 (both the amount and vertical profile), observation geometry (θ_0, θ, and ϕ), and the pressure and reflectivity (R) of the underlying clouds or surface. It can be parameterized and interpolated from the Jacobian lookup table following a similar approach as in eq. 3:

$$m(z) = \left[\frac{\partial I_0(\theta_0,\theta)}{\partial \tau_{so2}(z)} + \frac{\partial I_1(\theta_0,\theta)}{\partial \tau_{so2}(z)}\cos\phi + \frac{\partial I_2(\theta_0,\theta)}{\partial \tau_{so2}(z)}\cos 2\phi + \frac{R}{(1-RS_b)}\frac{\partial I_r(\theta_0,\theta)}{\partial \tau_{so2}(z)} + \frac{R^2 I_r(\theta_0,\theta)}{(1-RS_b)^2}\frac{\partial S_b}{\partial \tau_{so2}(z)}\right] \cdot I^{-1}. \quad (5)$$

For each of the 46 ozone climatology profiles derived by Labow et al. [2015] from ozone sonde and Microwave Limb Sounder (MLS) measurements, we ran VLIDORT simulations to build a Jacobian lookup table with dimensions of 6 × 8 × 8 × 72 × 801. The first three dimensions (6, 8, 8) correspond to the six nodes for the underlying surface/cloud pressure (240-1013 hPa), eight nodes for SZA (0-81°), and eight nodes for VZA (0-80°), respectively. The last two dimensions are used to store vertically (72 layers, 0.01-1013.25 hPa) and spectrally (305-345 nm at 0.05 nm resolution) resolved parameters for Jacobians. Unlike the lookup tables for volcanic SO_2, the tables for anthropogenic SO_2 do not include a dimension for SO_2 amounts. It is assumed that under the vast majority of conditions, the loading of anthropogenic SO_2 is relatively small and would have negligible effects on the scattering weight.

To account for the effects of clouds on OMI sensitivity to SO_2 for partially cloudy scenes, we employ the independent pixel approximation (IPA) approach commonly used in UV/VIS trace gas retrievals [*e.g.*, *Ahmad et al.*, 2004; *Koelemeijer et al.*, 2001; *Martin et al.*, 2002; *Seftor et al.*, 1994]. For each OMI scene, we use effective cloud fraction (f_c) and optical centroid cloud pressure (OCP) from the OMCLDRR product [*Joiner and Vasilkov*, 2006], derived at wavelengths near the fitting window for SO_2, to estimate scattering weight under overcast conditions ($m_{cld}(z)$, $f_c = 1$). We also use the same surface reflectivity and surface pressure as used in OMCLDRR retrievals to estimate scattering weight ($m_{cld}(z)$) under cloud free conditions ($m_{clr}(z)$, $f_c = 0$). We weigh the cloudy and cloud-free scattering weights with the cloud radiance fraction (CRF) within the SO_2 fitting window:

$$m(z) = m_{cld}(z) CRF + m_{clr}(z)(1 - CRF), \quad (6)$$
$$CRF = f_c \frac{I_{cld}}{I_{meas}}, \quad (7)$$

where I_{cld} and I_{meas} are calculated radiances with $f_c = 1$ and measured radiances at TOA, respectively, and z is the altitude above the sea level (represented by the pressure of each layer in the scattering weight and *a priori* profiles).

The column SO_2 Jacobian and air mass factor (AMF) represent the sensitivity of TOA radiances to change in SO_2 loading in the entire atmospheric column. The two parameters are linked through the absorption cross sections of SO_2 and can both be calculated from the scattering weight and the *a priori* profile of SO_2, $n_{SO2}(z)$. For example: :

$$AMF = \int_0^{TOA} m(z)\, n_{SO2}(z) dz. \quad (8)$$

Note that $n_{SO2}(z)$ is normalized against the VCD and represents the fraction of SO_2 molecules contributed by layer z to the overall SO_2 molecules within the entire atmospheric column.

In OMSO2 version 2.0 anthropogenic SO_2 retrievals, we estimate the column SO_2 Jacobians at all wavelengths between 310.5 and 340 nm and use the spectral Jacobians to estimate the SO_2 VCD based on eq. 1. In addition, we also provide an estimate of the slant column density (SCD) by fitting the PCs and the cross sections of SO_2 to the measured radiances. The SCD can be converted to VCD using *AMF*:

$$VCD = \frac{SCD}{AMF}. \tag{9}$$

In OMSO2 version 2.0, SCD is provided along with the scattering weight profile at 313 nm for each OMI scene so that data users have the flexibility to use their own *a priori* profiles to estimate AMF (eq. 8) and VCD (eq. 9).

4) Updated *a priori* profiles based on model simulations. For *a priori* profiles ($n_{SO2}(z)$ in eq. 8), we use GEOS-5 (Goddard Earth Observing System, Version 5) global model simulations (72 vertical layers, 0.5° latitude by 0.667° longitude horizonal resolution) that cover the period of 2004-2014. The output from GEOS-5 was sampled at the model's original spatial resolution and at the OMI overpass time and normalized against the total model simulated VCD within each grid cell to produce monthly normalized profiles. For each month of the year, these normalized profiles are then averaged over the entire simulation period to generate monthly climatology profiles that are used as *a priori* in the retrievals of ColumnAmountSO2. The *a priori* profile shape of SO_2 ($n_{SO2}(z)$) used for each OMI scene is provided in the data field GEOS5LayerWeight in OMSO2 product version 2.0.

In addition to retrievals with GEOS-5 based *a priori* profiles, we also produce SO_2 VCD retrievals using the same Jacobian lookup table, but different *a priori* profiles that assume constant SO_2 mixing ratio within the lowest 1 km of the atmosphere and negligible SO_2 above 1 km. SO_2 VCDs (ColumnAmountSO2_PBL) retrieved this way are independent from GEOS-5 simulations, and could provide useful observational constraints when the emission inventory used in model simulations is inaccurate or incomplete (for example, in the case of some large sources being absent or significantly underestimated in the inventory). ColumnAmountSO2_PBL is retrieved for each OMI scene with CRF < 0.5, and the corresponding *a priori* profile shape (data field name: PBLLayerWeight) is also provided for each scene in OMSO2 version 2.0.

5) Improved retrievals over snow/ice covered areas. The highly reflective snow/ice covered surfaces enhance OMI sensitivity to SO_2 in the lower atmosphere, but retrievals over these areas in the previous OMI PBL SO_2 dataset (ColumnAmountSO2_PBL in OMSO2 version 1.3 and before) are generally biased high as the snow/ice effects are not properly accounted for. In OMSO2 version 2.0, we use OMCLDRR product [*Vasilkov et al.*, 2010] and snow/ice flag in OMI L1B data (OML1BRUG) to identify scenes that are cloud-free and covered by snow/ice. For these scenes, we assign an algorithm flag value of 1 (i.e., AlgorithmFlag_SnowIce = 1), assume cloud fraction of zero and use measured Lambertian equivalent reflectivity (LER) for the whole scene in Jacobian calculations and SO_2 retrievals. For some snow/ice scenes, unambiguous cloud detection is not possible. While we also use measured LER from these scenes in SO_2 retrievals, we assign

an algorithm flag value of 2 (i.e., AlgorithmFlag_SnowIce = 2) so that they can be excluded from data analysis.

6) Multiple estimates of SO_2 abundances for different applications. As a summary of what has been described above, in OMSO2 version 2.0, three estimates for SO_2 abundances are now provided for different applications and studies related to air quality:

- **SlantColumnAmountSO2** (SCD) is the slant column density of SO_2 produced by fitting PCs and SO_2 cross sections to OMI measured radiances. SCD is introduced as a continuity dataset for the original PCA-based OMI PBL SO_2 data (OMSO2 version 1.2 and 1.3) using fixed SO_2 Jacobians (see section 3.2), especially when it is converted to VCD using a fixed AMF of 0.36.

- **ColumnAmountSO2** is an estimate of SO_2 VCD produced with SO_2 Jacobians from a more extensive lookup table and GEOS-5 based monthly *a priori* profiles. This dataset is recommended for most applications and analyses related to anthropogenic SO_2 pollution.

- **ColumnAmountSO2_PBL** is an estimate of SO_2 VCD produced with SO_2 Jacobians from the same lookup table as in ColumnAmountSO2, but different *a priori* profiles that assume constant SO_2 mixing ratio below 1 km and negligible SO_2 above 1 km. ColumnAmountSO2_PBL is independent from model simulations and the dataset is recommend for top-down estimates of emission sources. ColumnAmountSO2_PBL is only provided for scenes with cloud radiance fraction (CRF) < 0.5.

7) Additional auxiliary parameters for data users. In addition to retrieved SO_2 VCDs, we also provide new auxiliary datasets in OMSO2 version 2.0 to facilitate data analysis. They include:

AlgorithmFlag_SnowIce: value of 1 indicates that the scene is covered by snow/ice and likely cloud-free, and that measured LER from the scene has been used in Jacobian calculations; value of 2 indicates that the scene is covered by snow/ice but whether it is cloud-free is uncertain, and that the scene should be excluded from data analysis; value of 3 indicates that the scene is covered by snow/ice and likely covered by clouds, and that no special treatment has been applied to the scene. Value of 0 indicates that the scene is not covered by snow/ice and no special treatment has been applied to the scene.

Scattering weight: scattering weight (or vertically resolved box-AMF) at 313 nm at 72 vertical layers (as defined by LayerBottomPressure and a TOA pressure of 0.01 hPa) is interpolated from the Jacobian lookup table and provided for each scene. Users can use their own *a priori* profile to derive AMF and the corresponding SO_2 VCD for each OMI scene following Eq. 8 and 9.

LayerBottomPressure: pressure at the lower edge of each of the 72 vertical layers for which scattering weight is provided. User supplied *a priori* profile needs to be interpolated to these 72 layers. The pressure at the top of the highest vertical layer or the top of the atmosphere is 0.01 hPa.

GEOS5LayerWeight: GEOS-5 based *a priori* profile used in the retrieval of ColumnAmountSO2 for each scene. The profile is normalized for each scene (i.e., the sum of GEOS5LayerWeight from all 72 layers is one) and the value represents the fraction of SO_2 molecules each layer contributes to the entire atmospheric column.

PBLLayerWeight: *a priori* profile used in the retrieval of ColumnAmountSO2_PBL for each scene. It is also normalized for each scene (i.e., the sum of PBLLayerWeight from all 72 layers is one) and the value represents the fraction of SO_2 molecules each layer contributes to the entire atmospheric column.

CloudRadianceFraction: cloud radiance fraction (CRF) at 313 nm is provided. It is recommended that only scenes with CRF < 0.5 are used for air quality applications.

Flag_RowAnomaly: a flag that indicates whether a scene is potentially affected by the OMI row anomaly. Scenes with Flag_RowAnomally = 1 should be excluded from data analysis.

Flag_SAA: a flag that indicates whether a scene is potentially affected by the south Atlantic anomaly (SAA). Scenes with Flag_SAA = 1 should be excluded from data analysis. Note that Flag_SAA only flags scenes that are strongly affected by SAA. Some scenes with Flag_SAA = 0 may still be affected by SAA, although probably to a lesser extent.

4. Data Quality Assessment

Errors in OMI SO_2 data can arise from both the input radiance data and the SO_2 Jacobians used in retrievals. The resulted errors are best described as pseudo-random (i.e. having different systematic and random components depending on spatial and temporal scales) Gaussian-like distribution with a nominal mean of zero.

Below are data quality assessments for each SO_2 dataset included in OMSO2 product. For all datasets the noise increases with increasing solar zenith angle at high latitudes and in the region of South Atlantic Anomaly. **OMI scenes with Flag_RowAnomaly =1 should not be used**.

SlantColumnAmountSO2 (new in OMSO2 version 2.0): as an estimate of retrieval noise, the typical spatial standard deviation (σ) for instantaneous field of view (IFOV) is ~0.2 DU or smaller over the presumably SO_2-free remote Pacific when SZA is less than 50° and increases to ~0.3-0.4 DU at SZAs of 50-70°. There is a relatively small cross-track dependence as well, with OMI rows near the edge of the swath generally having greater noise. If we convert the SCDs to VCDs using a fixed AMF of 0.36, the 1-σ noise of VCDs will be ~0.5 DU at relatively small SZAs. This is comparable with the assessed noise level for our previous version of OMI PBL SO_2 dataset in OMSO2 versions 1.2 and 1.3, which effectively uses the same AMF.

The mean SCD over the remote Pacific is typically ~0.02-0.03 DU for SZAs up to 70°, suggesting overall very small systematic biases over background areas. The cross-track dependence of the mean SCD is generally small, and the maximal row-to-row difference is typically less than 0.05 DU. The spread in the mean SCDs between different rows for SZA > 70° may reach up to 0.1 DU, indicating larger cross-track dependence at large SZAs.

When using SlantColumnAmountSO2 as a continuity product for our previous PBL SO_2 dataset (ColumnAmountSO2_PBL in OMSO2 versions 1.2 and 1.3), data users should expect mostly consistent results with the previous version. As with our previous version OMSO2 product, **data users are also advised to use caution when analyzing data from the edges of the OMI swath (rows 0 and 59, 0-based), as they tend to have greater noise. For best data quality, use data from scenes near the center of the swath (rows 4-54, 0-based) with SZA < 65°. Retrievals for OMI scenes from the descending node of the Aura satellite should not be used. SlantColumnAmountSO2 retrievals for OMI scenes with CRF > 0.3 or non-zero AlgorithmFlag_SnowIce should not be used** (see below for instructions on ColumnAmountSO2 and ColumnAmountSO2_PBL).

ColumnAmountSO2 (new in OMSO2 version 2.0): The use of new SO_2 Jacobian lookup table and GEOS-5 *a priori* profiles helps to significantly reduce the noise of ColumnAmountSO2 particularly over relatively clean background areas. Note that the noise level of ColumnAmountSO2 depends on both the spectral fitting (i.e., SCDs) and the column Jacobians/AMFs. The 1-σ noise of ColumnAmountSO2 is typically < 0.1 DU over remote oceanic areas, as SO_2 is commonly more abundant at higher altitudes in *a priori* profiles used for those areas, leading to larger AMFs. Over continental areas, *a priori* profiles tends to have more SO_2 at lower altitudes, and the AMFs are normally smaller, often leading to greater noise in ColumnAmountSO2 particularly over polluted areas at greater SZAs and/or VZAs.

The SO_2 retrieval accuracy also depends on the error in the SO_2 Jacobians. There are multiple sources that may cause error in the Jacobians/AMFs. These include: 1) errors from forward radiative transfer model assumptions and the table lookup interpolation scheme for the scattering weight, which are typically < 10-15%, but can become larger at higher SZAs and/or VZAs; 2) errors in *a priori* profiles. Comparisons with limited available aircraft measurements have shown that inaccurate monthly *a priori* profiles may lead to up to 40% of error over the polluted northeastern China; 3) errors in input data such as cloud pressure, cloud fraction, and surface reflectivity. Depending on the vertical distribution of the cloud layer and the SO_2 layer, a typical ~50-100 hPa uncertainty in cloud pressure can cause ~50% error in ColumnAmountSO2 over polluted areas (where SO_2 is mostly near the surface) and have negligible effects over background regions where SO_2 is typically more elevated. A typical uncertainty of ~0.05 in cloud fraction leads to an error of ~5-10% in SO_2, whereas an error of ~0.01 in surface reflectivity leads to ~10% error in SO_2; 4) Aerosol effects. The effects of aerosols on SO_2 Jacobians are not explicitly accounted for in the current product. For non-absorbing aerosols such as sulfate, the error in SO_2 VCDs resulted from the lack of explicit aerosol treatment is typically 20% or less, but for strongly absorbing aerosols such as smoke, dust or ash, the error can be as large as 50%. Assuming all these sources of errors are independent, the overall relative uncertainty of AMF/Column Jacobians would be ~30-100%.

As with SlantColumnAmoutnSO2, when using ColumnAmoutnSO2, it is recommended that data users exclude scenes in the outermost rows (0 and 59, 0-based), or with large cloud fraction (CRF > 0.5), or at large solar zenith angle (> 70°). **For best data quality, use data from scenes near the center of the swath (rows 4-54, 0-based) with SZA < 65°, CRF < 0.3, and AMF > 0.3. Retrievals for OMI scenes from the descending node of the Aura satellite should not be used. Scenes with AlgorithmFlag_SnowIce = 2 should not be used.**

ColumnAmountSO2_PBL (updated in OMSO2 version 2.0): starting from version 2.0, ColumnAmountSO2_PBL retrievals use the same Jacobian lookup table as in ColumnAmountSO2. With *a priori* profiles assuming essentially all SO_2 in the lowest 1 km of the atmosphere, the noise in ColumnAmountSO2_PBL is generally greater than **ColumnAmountSO2, especially over remote background areas. Therefore, ColumnAmountSO2_PBL should only be used for studies that aim to estimate the emissions from known or detected sources.** ColumnAmountSO2_PBL is only provided for CRF < 0.5, and for best data quality, the same data screening criteria as those for ColumnAmountSO2 are recommended (i.e., **use rows 4-54, 0-based, SZA < 65°, CRF < 0.3, AMF > 0.3, and AlrogithmFlag_SnowIce ≠ 2**),

ColumnAmountSO2_TRL: Due to increased sensitivity to elevated SO_2, the pixel-level 1-σ noise in TRL data is estimated at ~0.2 DU under optimal observational conditions in the tropics. The noise is about 0.3 DU for high latitudes. The data can be used for cloudy, clear and mixed scenes as well as for elevated terrain, but will overestimate SO_2 VCDs if the plume is present at altitudes higher than 3 km. Since the noise level of the PCA TRL SO_2 is about a factor of two smaller than the previous Linear Fitting (LF) TRL SO_2, we recommend that the TRL retrievals be used for estimating emissions from degassing volcanoes.

ColumnAmountSO2_TRM: The standard deviation of TRM retrievals over background areas is about 0.1 DU in the tropics and about 0.15 DU at high latitudes, again about a factor of two smaller than the previous LF TRM retrievals. Like the TRL data, the TRM data can be used for various sky conditions. The TRM data can be used for investigating SO_2 plumes from moderate eruptions.

ColumnAmountSO2_TRU data **(new in OMSO2 version 2.0)** are intended for use for explosive volcanic eruptions where SO_2 is injected into the upper troposphere. The standard deviation over background areas is around 0.1 DU for all latitudes for TRU data.

ColumnAmountSO2_STL data are intended for use for explosive volcanic eruptions where SO_2 is injected into the upper troposphere or lower stratosphere (UTLS). The standard deviation over background areas is around 0.1 DU for all latitudes for STL data.

Unlike the previous LF algorithm that has large negative bias for high SO_2 loading cases (> 100 DU), the PCA volcanic SO_2 algorithm has greatly reduced this error and compares well with other retrieval methods that also utilize the full spectral content of UV hyperspectral measurements. For example, for the Kasatochi eruption in August 2008, the PCA-estimated total SO_2 loading is ~1700 kt, about a factor of two higher than the estimate based on LF retrievals and generally in good agreement with the offline OMI iterative spectral fitting (ISF) retrievals and GOME-2 optimal estimation retrievals. For the Sierra Negra eruption in October 2005, the PCA algorithm yields a max SO_2 VCD of over 1100 DU. This agrees well with the offline OMI ISF retrievals and is several times greater than the max SO_2 from the previous LF algorithm. It should be noted that the PCA algorithm might still underestimate SO_2 loading for very dense plumes immediately after large eruptions. Additionally, some eruptions also emit large amounts of ash into the atmosphere that may interfere with SO_2 retrievals. This interference may not necessarily be appropriately accounted for using the SLER approach and may cause errors in the retrieved SO_2.

Volcanic SO_2 data from all rows of the OMI, with the exception of rows affected by the row anomaly, can be used. I**t is best to use retrievals from scenes with SZA < 70°, and retrievals for OMI scenes from the descending node of the Aura satellite should be excluded.** When estimating the SO_2 loading from a volcanic plume within a given domain, it is recommended that only OMI scenes or pixels exceeding a certain threshold (e.g., 1 DU) be included in the calculation. This helps to filter out occasional negative retrieval noise. One caveat of this approach, when applied to a relatively small domain, is that the estimated total SO_2 mass depends on the selected threshold, particularly for those relatively small eruptions. Alternatively, data users may also select a large negative threshold (e.g. SO_2 > -10 DU) and a relatively large geographical domain (including both plume and background region) for SO_2 mass calculation.

5. Product Description

The OMSO2 product is written as a collection of HDF-EOS5 swath files. The data files are available from Goddard Earth sciences Data and Information Services Center (GES DISC) web site (Li et al., 2019; https://doi.org/10.5067/Aura/OMI/DATA2022).

A product file, also called a data granule, contains SO_2 and associated information retrieved from each OMI scene from the sun-lit portion of an Aura orbit. The data are ordered in time sequence. The information provided in these files includes: latitude, longitude, solar zenith angle, and independent estimates of the SO_2 vertical columns, as a well as a number of ancillary parameters that provide information to assess data quality. Six values of SO_2 VCD are provided corresponding to different assumed vertical profiles.

Independent information is needed to decide which value is most applicable. For a complete list of the parameters, please read the OMSO2 file Specification.

For general assistance with data archive, please contact GES DISC. For questions and comments related to the OMSO2 algorithm and data quality please contact Can Li (Can.Li@nasa.gov), who has the overall responsibility for this product, with copies to Nickolay Krotkov (Nickolay.A.Krotkov@nasa.gov).

The subsets of OMSO2 data over many ground stations and along Aura validation aircraft flights paths are also available through the Aura Validation Data Center (AVDC) web site under data/Aura.

6. References

Ahmad, Z., P. K. Bhartia, P. K., and N. Krotkov (2004), Spectral properties of backscattered UV radiation in cloudy atmospheres, *J. Geophys. Res.*, 109, D01201, doi:10.1029/2003JD003395.

Bhartia, P. K. and C. W. Wellemeyer (2002), OMI TOMS-V8 Total O3 Algorithm, *Algorithm Theoretical Baseline Document: OMI Ozone Products*, edited by P. K. Bhartia, vol. II, ATBD-OMI-02, version 2.0. Available: http://eospso.gsfc.nasa.gov/eos_homepage/for_scientists/atbd/docs/OMI/ATBD-OMI-02.pdf

Bogumil, K., J. Orphal, T. Homann, S. Voigt, P. Spietz, O. C. Fleischmann, A. Vogel, M. Hartmann, H. Kromminga, H. Bovensmann, J. Frerick, J. P. Burrows (2003), Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230-2380 nm region, Journal of Photochemistry and Photobiology, A: Chemistry, 157 , 167-184.

Fisher, B. L., N. A. Krotkov, P. K. Bhartia, C. Li, S. A. Carn, E. Hughes, and P. J. T. Leonard (2019), A new discrete wavelength backscattered ultraviolet algorithm for consistent volcanic SO_2 retrievals from multiple satellite missions, *Atmos. Meas. Tech.*, 12, 5137–5153, https://doi.org/10.5194/amt-12-5137-2019.

Joiner, J., L. Guanter, R. Lindstrot, M. Voigt, A. P. Vasilkov, E. M. Middleton, K. F. Huemmrich, Y. Yoshida, and C. Frankenberg (2013), Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2, *Atmos. Meas. Tech.*, 6, 2803-2823, doi:10.5194/amt-6-2803-2013.

Koelemeijer, R. B. A., P. Stammes, J. W. Hovenier, and J. F. deHaan (2001), A fast method for retrieval of cloud parameters using oxygen A band measurements from the

Global Ozone Monitoring Experiment, *J. Geophys. Res.*, 106(D4), 3475–3490, doi:10.1029/2000JD900657.

Krotkov, N. A., et al. (2008), Validation of SO_2 retrievals from the Ozone Monitoring Instrument over NE China, *J. Geophys. Res.*, 113, D16S40, doi:10.1029/2007JD008818.

Krotkov, N. A., S. A. Carn, A. J. Krueger, P. K. Bhartia, and K. Yang (2006). Band residual difference algorithm for retrieval of SO_2 from the Aura Ozone Monitoring Instrument (OMI). *IEEE Trans. Geosci. Remote Sensing*, 44(5), 1259-1266, doi:10.1109/TGRS.2005.861932, 2006.

Krueger, A. J., L. S. Walter, P. K. Bhartia, C. C. Schnetzler, N. A. Krotkov, I. Sprod, and G. J. S. Bluth (1995) Volcanic sulfur dioxide measurements from the total ozone mapping spectrometer instruments. *J. Geophys. Res.*, 100(D7), 14057-14076, 10.1029/95JD01222.

Labow, G. J., J. R. Ziemke, R. D. McPeters, D. P. Haffner, and P. K. Bhartia (2015), A total ozone-dependent ozone profile climatology based on ozone sondes and Aura MLS data, *J. Geophys. Res. Atmos.*,120, 2537–2545, doi:10.1002/2014JD022634.

Li, C., J. Joiner, N. A. Krotkov, and P. K. Bhartia (2013), A fast and sensitive new satellite SO_2 retrieval algorithm based on principal component analysis: Application to the Ozone Monitoring Instrument, *Geophys. Res. Lett.*, 40, doi:10.1002/2013GL058134.

Li, C., N. A. Krotkov, S. Carn, Y. Zhang, R. J. D. Spurr, and J. Joiner (2017), New generation NASA Aura Ozone Monitoring Instrument (OMI) volcanic SO_2 dataset: Algorithm description, initial results, and continuation with the Suomi-NPP Ozone Mapping and Profiler Suite (OMPS), *Atmos. Meas. Tech.*, 10, 445–458.

Li, C., N. A. Krotkov, P. Leonard, and J. Joiner (2019), OMI/Aura Sulphur Dioxide (SO_2) Total Column 1-orbit L2 Swath 13x24 km V003, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) (2019), Accessed: Nov. 10, 2019, 10.5067/Aura/OMI/DATA2022.

Martin, R. V., et al. (2002), An improved retrieval of tropospheric nitrogen dioxide from GOME, *J. Geophys. Res.*, 107(D20), 4437, doi:10.1029/2001JD001027.

Seftor, C. J., S. L. Taylor, C. G. Wellemeyer, and R. D. McPeters (1994), Effect of partially- clouded scenes on the determination of ozone, Ozone in the Troposphere and Stratosphere, edited by R. D. Hudson, NASA Conference Publication 3266, 919-922.

Stammes, P., M. Sneep, J. F. de Haan, J. P. Veefkind, P. Wang, and P. F. Levelt (2008), Effective cloud fractions from theOzone Monitoring Instrument: Theoretical framework and validation, *J. Geophys. Res.*, 113, D16S38, doi:10.1029/2007JD008820.

Yang, K., N. Krotkov, A. Krueger, S. Carn, P. K. Bhartia, and P. Levelt (2007), Retrieval of Large Volcanic SO_2 columns from the Aura Ozone Monitoring Instrument (OMI): Comparisons and Limitations, *J. Geophys. Res.*, 112, D24S43, doi:10.1029/2007JD008825

Appendix A. Changes in data fields between OMSO2 Version 1.3 and Version 2.0

Table A.1 Changes in Data fields between OMSO2 Version 1.3 and Version 2.0

Data Field Name in OMSO2 Version 1.3	Changes Made in OMSO2 Version 2.0
AlgorithmFlag_PBL	Removed
AlgorithmFlag_STL	Removed
AlgorithmFlag_TRL	Removed
AlgorithmFlag_TRM	Removed
ChiSquareLfit	Removed
ColumnAmountSO2_PBLbrd	Removed
ColumnAmountSO2_STLbrd	Removed
ColumnAmountSO2_TRMbrd	Removed
LayerEfficiency	Removed
QualityFlags_PBL	Removed
QualityFlags_STL	Removed
QualityFlags_TRL	Removed
QualityFlags_TRM	Removed
Reflectivity331	Replaced with Reflectivity342
Residual	Removed
ResidualAdjustment	Removed
Rlambda1st	Removed
Rlambda2nd	Removed
SO2indexP1	Removed
SO2indexP2	Removed
SO2indexP3	Removed
Wavelength	Removed
dN_dSO2_STL	Removed
dN_dSO2_TRL	Removed
dN_dSO2_TRM	Removed
deltaO3	Removed
deltaRefl	Removed
fc	Removed
Data fields that do not exist in OMSO2 Version 1.3 but are added in OMSO2 Version 2.0	
	AlgorithmFlag_SnowIce
	CloudFraction
	CloudRadianceFraction
	ColumnAmountSO2
	ColumnAmountSO2_TRU
	Flag_RowAnomaly
	Flag_SAA
	GEOS5LayerWeight
	LayerBottomPressure
	PBLLayerWeight
	ScatteringWeight

	SceneReflectivity354
	SlantColumnAmountSO2
	SurfaceReflectivity
	GroundPixelArea
	TiledCornerLatitude
	TiledCornerLongitude